Factors associated with private-public school performance: analysis of TALIS-PISA link data

Article (Accepted Version)

Factors Associated with Private-Public School Performance: Analysis of TALIS-PISA link data

Marcos Delprato[1],[2]
Centre for International Education, University of Sussex, Brighton BN1 9QQ UK Email: m.delprato@sussex.ac.uk

Amita Chudgar
College of Education, Michigan State University, 408, Erickson Hall, East Lansing, MI 48824, USA Email: amitac@msu.edu

Abstract

We use measures of competitive pressure, administrative autonomy and staffing practices to explain the private-public performance difference in Australia, Portugal and Spain using the TALIS-PISA dataset. We employ OLS regression and counterfactual decomposition analysis on matched sub-samples. These school factors do not explain the overall private-public performance gap in the three countries except at the higher-end of the distribution. In other words, these factors appear to benefit only the high-performers in private schools in Australia and Spain. The results point to the potential limits of adopting private school practices for improving learning across the performance distribution especially for low-performing students.

Keywords: Private schools; Systemic school factors; Learning gap; TALIS-PISA; Propensity Score; Counterfactuals; Decomposition.

---

[1] Corresponding author.
[2] The authors would like to acknowledge Jutaro Sakamoto at Michigan State University for research assistance and Dr. Susan Printy also at Michigan State for her guidance on some of the literature we refer to. All errors are our own.

1. Introduction

The debate on the efficiency of private versus public schools has a long history. However, an empirical understanding of the role of `systemic' features of educational institutions and outside organization factors (e.g., competition, autonomy) behind public and private schools' differential performance is still limited, particularly from an international perspective.

This paper focuses on understanding how systemic differences between private and public educational institutions (namely competitive pressure, administrative autonomy, staffing practices and accountability) can explain differences in students' performance (measured by maths, reading and science test scores) in three countries with diverse education systems: Australia, Portugal and Spain. For this purpose, we rely on the Teaching and Learning International Survey-Program for International Student Assessment (TALIS-PISA) which provides a unique set of student and school variables relevant for our study.

Few considerations make TALIS-PISA link the ideal dataset to answer our main research question. First, although in isolation PISA is missing the teacher level data and TALIS the student level data (Kaplan and McCarty, 2013), TALIS-PISA provides a linkage of the two key elements of learning outcomes within public and private schools. On the one hand, TALIS includes, among others information on how the work of teachers is recognised, reviewed, supported and rewarded and the role of school leaders (Austin et al., 2015; Sallies et al., 2016) and, on the other hand, characteristics of the students and schools are included in PISA (OECD, 2012a). Second, TALIS offers unique, detailed (e.g., teacher demographic characteristics, school leadership, teacher professional development, systems of feedback and appraisals) as well as internationally comparable information on learning environments and the working conditions in schools (Freeman et al., 2014).[1] Indeed, as our paper is an empirical cross-country study, having comparable reliable dataset on systemic differences between private and public educational institutions is critical.

A central issue in studying the private-public performance difference is the problem of identification arising from selection bias; students and families select public or private schools and are not placed in them randomly. Students from more affluent and educated families are more likely to be enrolled in private schools resulting in a systematically different social composition within private schools compared to public schools (Dronkers and Robert, 2008). This complicates isolating the so-called `private school effect' (Hanushek, 1986; Rivkin et al., 2005).

Looking at the complete characterization of the relationship between learning scores and private and public enrolment, beyond the students at the mean, is important too. Numerous studies look at effects at different performance quantiles to explain gaps in achievement for contrasting groups; for instance: gender differences (Robinson and Lubienski, 2011), private-public school differences (Oliveira et al., 2013), urban-rural differences (Lounkaew, 2013), and for instruction time in low achieving schools (Battistin and Meroni, 2016). A quantile approach to private-public learning gap is similarly relevant since it is linked to the effect of school systemic factors mediated by school quality (measured by students' performance quantile). It is also important methodologically since unobserved family inputs and ability are related to where a student's performance falls within the learning distribution.

After accounting for identification issues underlying the public private learning gap in these three countries, we investigate to what extent are systemic school factors important in understanding the differential performance of public and private schools. In other words, which systemic factors influence the private-public performance gap all else equal. We also ask if the effect of these systemic school factors vary by whether a student is a low or a high performer to understand more closely the potential mechanisms through which private schools may (or may not) outperform public schools.

To conduct this analysis, we proceed in three steps. First, we generate a matched sub-sample of comparable private and public students by using matching techniques. Second, we present estimates for each systemic school factor on this matched sub-sample. Finally, we use a counterfactual decomposition analysis on the matched sub-sample to distinguish how the association of systemic school factors with the private-public learning gap can be decomposed into the changes in the covariates and changes in the coefficients, putting particular emphasis on how the return to each systemic school factor changes at different points of the learning distribution. This goes beyond mean effects and provides information about the entire conditional distribution of student performance. On equity grounds, knowing whether there is heterogeneity in the effect of systemic factors by students' performance levels is vital in identifying which factors lead to larger gaps for the more disadvantaged/low performing public school students.

As an illustration of what insights can be gained through the counterfactual analysis assume, for instance, that school accountability is the systemic factor of interest. The counterfactual analysis based on matched sub-samples would tell us (after isolating differences in performance due to students' personal characteristics) whether private schools' students would perform worse than public schools' students if they had attended schools with the same average distribution of school accountability levels as public schools' students. Additionally, this analysis shows how this difference changes across the learning distribution. In other words, if accountability varies with performance, then the counterfactual analysis would help us to identify whether the initial positive effect of more accountability among low performers in private schools fades away for top performers because of diminishing returns to characteristics with scores. The counterfactual analysis helps to identify which systemic school factors contributes more to the learning gap and for which group of achievers. In addition, it allows us to divide the total effect of systemic factors into the effect of the coefficient and the effect of the level of the systemic school factor.

As far as we are aware, Oliveira et al. (2013) is the only study using counterfactual analysis to focus on the private-public learning gap in Brazil. Their analysis focuses purely on the impact of individual and regional characteristics on the school type learning gap but not on the role of school systemic factors. Our study also departs from their work because it is a cross-country study and we also employ matched sub-samples, thereby our conclusions are net of individual's student characteristics.[2]

The rest of the paper is organised as follows. Section 2 contains a brief review on the mechanisms linking the above mentioned systemic school factors (competitive pressure, administrative autonomy, staffing practices and accountability) with students' achievement, as well as a brief description of each country's education system. The data used is described in Section 3. In Section 4 we explain the empirical approach and results are presented in Section 5. A summary of the main findings and implications is included in Section 6.

2. A brief review of key systemic school factors and country background

The literature on public and private school performance is extensive and a range of overlapping, school-based mechanisms are discussed to explain their potential performance differences.

2.1 Importance of competitive pressures, administrative autonomy, and staffing practices

The importance of competitive pressure in improving school performance is perhaps one of the oldest and strongest arguments in favour of school choice (e.g., Friedman, 1955). The basic notion is that the presence of competition (versus a monopoly) will force all actors—public and private alike—to 'become more efficient in response to competition' (Hoxby 2001, p. 142). The arguments for greater innovation are also closely linked to the importance of introducing greater competition and thus choice in the education system. According to this argument, a successful education system will promote '…unleashing the productive potential already present in the schools and their personnel', whereas over-reliance on bureaucratic structures, rules, and regulations may serve to undermine these innovative tendencies of the schools and their staff (Chubb and Moe, 1990, p. 5). Indeed, the charter school phenomenon in the United States is predicated on the importance of innovation in charter success (e.g., Lubienski, 2004).

Accountability has been understood as a mechanism to improve the quality of education by making schools accountable for attaining high standard (Geijsel et al., 2010). The combination of more autonomy and accountability is expected to improve schooling by promoting effectiveness and efficiency, and by binding education professionals (teachers and other staff members) to educational purposes and goals (Hooge et al., 2012). Increased school autonomy often goes together with stronger accountability demands: schools that are granted more decision-making power are required to account for these decisions and their impact.

Therefore, school governance and school autonomy −that is, how education responsibilities are managed and resources distributed− are concepts which permeates throughout the learning process impacting on teaching and learning. Many education systems (including those in Spain and Australia) have increased the role of schools in resource allocation and curriculum (Woessmann, 2007). This, of course, assumes both strong leadership and highly qualified teachers who can identify learning needs and at the same time are able to act upon these needs through evaluations and accountability systems without being overloaded (Hanushek et al., 2013). Those who support private schooling stress the importance of better accountability leading to more responsive private school. In a cross-country study for eight countries, Bloom et al. (2015) finds that the better performance of autonomous schools is not linked to larger autonomy but rather how autonomy is used and influenced by better school management −be it through having strong accountability of principals to an external governing body and exercising strong leadership through a coherent long-term strategy.[3]

A direct reflection of greater administrative autonomy is the ways in which school staffing decisions are handled. If school administrators have greater autonomy in decision-making, then it may reflect in their ability to both hire appropriate staff and to sanction or even fire such staff if their performance is found inadequate. Researchers have argued that greater reliance on school choice accompanied by greater autonomy in school functioning will lead to changes in the types of teachers that are hired and retained in such schools. And, overall, such a change will result in the hiring and retention of stronger teachers (e.g., Hoxby, 2002; Podgursky, 2006).

Another argument comes from the educational administration literature. The authors of this literature argue about the central role of school leaders in setting direction and developing people and organization for improving student performance. And that this in turn is associated with successful performance of their schools (e.g., Leithwood and Riehl, 2003). Presumably, school leaders with greater autonomy may be better able to not just sanction, but also support their staff and their organization.

Competition, autonomy, and the resultant ability to manage school staffing each suggest mechanisms through which public and private school performance may partly be shaped. If the differences along these criteria are vast between these schools, then they may also explain the variation in their performance. While scholars have extensively studied the importance of each of these factors in understanding school performance in general, to our knowledge, these factors have not been used systematically and cross-nationally to explain the private-public performance gap as we propose to do in this paper.

A key challenge facing such research is the difficulty of obtaining comparable data on public and private school competition, autonomy, and staffing along with student performance. This is a challenge we are able to overcome with the TALIS-PISA link data. In particular, TALIS contains rich information on the working conditions of teachers and the learning environments in schools. Before we introduce our data and empirical approach to address this question, we briefly discuss some relevant details about our three study countries.

2.2 Country background

The three countries studied here (Australia, Spain, and Portugal) exhibit important similarities and differences in their education systems. Australia performs higher than average on PISA compared to the other Organisation for Economic Co-operation and Development (OECD) countries and also enjoys greater than (OECD) average secondary and tertiary attainment.

Both Spain and Portugal have PISA performance as well as secondary and tertiary attainment that are at par or below par with the OECD average. Spain and Australia have above average private school participation; 18% of OECD students overall attend private schools while in Spain and Australia, this percentage is 40% and 34% respectively. In Portugal, private participation is 14%, which is relatively closer to the OECD average. In Australia, public contributions to education are below average, whereas in Spain and Portugal, the total public contributions are near the OECD average. Indeed, Australia has extensive school choice options; according to OECD reports from 2012 and 2013, 96% of students there attend a school that is facing some competitive pressure (OECD 2012b, 2013c).

In terms of centralisation and autonomy, the three systems also provide some variation. Education policy is guided nationally and delivered locally at state and territory levels by both government and private providers in Australia. Overall, schools there enjoy a fair amount of autonomy in decision making regarding instruction.

In Spain, national and regional ministries are important for educational decisions. The national level defines the overall framework, and so many of the schooling and funding decisions are made at the regional level with limited school autonomy (OECD 2014a).

Education policies are centralized at the national level with some recent degree of decentralisation in Portugal (OECD 2014b). Interestingly, school leadership roles in Portugal have only been professionalised since 2008 and overall, school leaders there report some of the lowest levels of autonomy with regard to curriculum and assessment among OECD countries.

A robust amount of literature has investigated private-public performance differences in each of these countries. In general, this literature finds somewhat consistent evidence for a positive private effect on student academic and behavioral outcomes (for example, Marks (2009) for Australia; Pereira (2011) for Portugal; and Mancebon and Muniz (2008) for Spain). However, this literature has paid limited attention to correcting for selection bias, and even less attention to explaining what factors may be driving these private-public gaps. In terms of selection bias, studies with robust socioeconomic controls do find a diminishing private effect (see for instance, Moran, Neri, and Rodgers (2014) for Australia; Rosado and Seabra (2015) for Portugal; and Crespo-Cebada, Pedraja-Chaparro, and Santin (2014) for Spain).

3. Data

This study is based on the TALIS-PISA link dataset. In the Teaching and Learning International Survey (TALIS) of 2013, eight countries opted to conduct the survey in schools that participated in PISA 2012 through the TALIS-PISA link option (OECD 2013a). In these countries, ‘a systematic equiprobable random sample of schools from PISA 2012’ (OECD

2013a, p:79) was identified for the TALIS-PISA link survey. The linked dataset combines the TALIS insights into the backgrounds, beliefs, and practices of teachers and school leaders with PISA information into backgrounds, beliefs, attitudes, and cognitive outcomes of students (Austin et al., 2015). On the contrary, in isolation, PISA only offers measures of student learning whereas TALIS only provides information about teachers' job- related attitudes (Kaplan and McCarty, 2013).

The TALIS-PISA link dataset follows a robust design by using the structure of the PISA 2012 sample of schools with original strata and setting a nominal sample for the TALIS-PISA link of 150 schools, interviewing an additional 20 teachers per schools and including schools if 50% of sampled teachers respond (OECD, 2013a, 2013b). The TALIS-PISA link dataset, nonetheless, contains some limitations (Ferrera and Izquierdo, 2016); that is, the two surveys of the joint TALIS-PISA link were not implemented at the same time (PISA in 2012 and TALIS in 2013). Still, since the timing of the two surveys' implementation is very close, the likelihood that teachers and school principals surveyed in TALIS may not be the same respondents who taught students evaluated in PISA, is therefore minimised.

We follow PISA's definition of private school which includes those that are managed locally, without regard to funding sources. Under this definition, private schools may or may not require parents to pay enrolment fees (OECD, 2011). Other PISA's studies distinguish between private independent and private government dependent schools (Dronkers, J., & Robert, 2008) while other studies don't make this distinction (Vandenberghe and Robin, 2004). We follow the latter approach because in our final sample all school falls into the private government dependent category as they receive between 66% and 88% of government funding.

In the current analysis, we excluded Finland, Latvia, Romania, and Singapore because the proportion of student attending private schools was below 2% or zero in those countries. We also dropped Mexico where the TALIS-PISA linked data reflected did not lead to a similar private-public relationship as the complete PISA data. This left us with the three countries included in our analysis: Australia, Portugal, and Spain. It is important to note that, in these countries, the distribution of students attending private schools remains nearly the same for the PISA and TALIS-PISA link datasets (39%–43% for Australia, 8%–9% for Portugal, and 38%–34% for Spain for PISA and TALIS, respectively) and the impact of school type on achievement is qualitatively similar in both PISA and TALIS-PISA link datasets. This highlights the representativeness of our sample[4], with the additional weighting in the TALIS-PISA link to account for sub-sampling within the original sample of PISA schools (OECD, 2013).

Tables 1 and 2 display summary statistics for the key variables used in this study. Table 1 includes the three performance outcomes and a set of student background controls we use for propensity score matching. Table 2 lists the school factors that we investigate to understand the private-public performance difference. These include the measure of principal-reported competitive pressure; principal-reported autonomy in staffing, budgeting, and instruction decisions, and principal-reported and teacher-reported sanctions and support offered to teachers as measures of staffing practices.

[Tables 1 and 2 here]

4. Empirical methods

The main objective of the paper is to examine the role of systemic school factors on the private-public achievement gap. We are also interested in allowing for heterogeneity in the response of systemic factors by students' learning levels. In other words, we ask whether a systematic factor such as competition would help explain private school performance, and if the presence of competition differs for high- and low-achieving students. For this we apply a counterfactual analysis on a comparable sub-sample. Through matching we obtain a private and public school students sub-sample with similar covariates[5] and then, based on this comparable sub-sample, we estimate the counterfactual (that is, the achievement difference that would have prevailed if private schools' students had the same returns to systemic characteristics that public schools' students have) decomposing the private-public learning scores differential across the learning distribution.

Our aim is find out whether students' achievement gap would have persisted if public and private schools had the same level of accountability, competition and staff policies, dividing the total difference by the impact of coefficients and characteristics in the spirit of the Oaxaca decomposition (Oaxaca, 1973) based on decomposition techniques (Albrecht et al., 2009; Chernozhukov et al., 2013; Machado and Mata, 2005; Melly; 2006).[6] As a robustness analysis, we also examine the role of unobservables on the private school effect using instrumental variables within a quantile framework (see Appendix B for details).

4.1 OLS regression and construction of matched samples

Our first step to disentangle the net private school effect is to control for an array of student covariates through OLS. For each country, we estimate the following OLS regression,

$$Y_i = \gamma_0 + \beta S_i + \gamma_1 X_i + \varepsilon_i \tag{1}$$

where Y_i denotes the test scores (maths, reading, and science literacy) for student i ($i = 1, \dots, N$), S_i is a dummy for school type (equal to 1 for private and 0 for public enrolment), and X_i is a set of students and family exogenous characteristics (from Table 1).

Next, to address selection bias introduced by school choice based on observables, we employ propensity score matching (Rosenbaum and Rubin, 1983) which yields more robust estimates for the private school effect. Matching estimators allow us to derive the counterfactual of the treated (the outcome that a private school student would have had if he had attended a public school) using information on untreated students with the same observable characteristics as the treated individuals (Dehejia and Wahba 1999).

More precisely, we estimate the average treatment on the treated (ATT) (i.e., those attending private school) as $\tau_{ATT} = E[Y(1) - Y(0)|X, S = 1]$, where $Y(1)$ and $Y(0)$ are the outcome of students attending private and public schools, respectively, and X denotes the set of observed covariates used to calculate the propensity score (the probability for each student of attending private school) which is then used to impute the missing outcome by finding other individuals in the data whose propensity scores are similar, but who were not exposed to the treatment. The estimator of the ATT relies on two key assumptions: unconfoundedness and overlap.[7] We run the propensity score matching analysis using nearest neighbour matching (without replacement) and use it to derive matched subsamples.[8]

4.2 OLS and counterfactual decomposition analysis based on matched subsample

In a second step, we employ OLS and a counterfactual decomposition analysis for matched subsamples of comparable private and public students in observed covariates. This matched sub-sample contains fewer observations than the original sample (i.e., $N_m < N$ as unmatched comparison units are discarded). For the average student, OLS provides an initial answer on the role played by systemic school factors on the private-public learning gap for the matched subsamples of N_m students,

$$Y_{i\prime} = \gamma_0 + \beta S_{i\prime} + \gamma_1 W_{i\prime} + \varepsilon_{i\prime} \qquad (2)$$

where $Y_{i\prime}$ is the outcome for matched student i' ($i' = 1, \dots, N_m$) and $W_{i\prime}$ denotes the school variables (autonomy, competition and staff policies).

Additionally, in the counterfactual decomposition analysis, learning differences at each point of learning distribution are decomposed into differences due to different students' performance between public and private schools with the same (systemic) W characteristics and a composition effect, arising due to public and private school difference on W characteristics. Formally, let F be the distribution function and F^{-1} the inverse function, and

0 and 1 denote the population of public and private students, respectively, so that $F_{(0|0)}$ and $F_{(1|1)}$ represent observed distribution of scores for public and private students, and let $F_{(0|1)}$ (the counterfactual) represents the distribution function of scores that would have prevailed for private school students had they faced public students' learning profile $F_{(0|0)}$. The total difference on the observed learning score quantile function (F^{-1}) between public and private students can be decomposed as:

$$F^{-1}_{Y(1|1)} - F^{-1}_{Y(0|0)} = \left[F^{-1}_{Y(1|1)} - F^{-1}_{Y(0|1)}\right] - \left[F^{-1}_{Y(0|1)} - F^{-1}_{Y(0|0)}\right] \qquad (3)$$

where the first term in brackets is due to differences in the coefficient effect for a given level of systemic school factors and the second term is the composition effect due to difference in the level of systemic school factors. One could think of the first term as an efficiency component measuring how a school characteristic translates into larger achievement and the second term as an aggregate effect linking the level of the school characteristic to achievement.

We look at the effect of coefficients and covariates independently (i.e., the two terms of equation 3) for the whole group of school systemic factors, while emphasis is placed on the effect of the level or aggregated value for each school systemic factor (the second term of equation 3).[9] Term 2 of equation (3) would inform us what a student attending a private school would have achieved if he/she had attended a school with same level of say, competition, of a student attending a public school.

5. Results

5.1. Summary statistics

Summary statistics for outcomes and explanatory variables are included in Table 1. This table shows that learning outcomes are consistently higher for students attending private schools, with private-public gaps ranging between 33–46 points in maths, reading, and science. Table 1 also shows that private school students are better-off compared to public students in terms of their observed characteristics. That is, they live in households with greater wealth as well as with more books and higher parental education. Students attending public schools are more likely to be from language minorities than private school students and more prone to school absence. Yet distinctions between public and private schools are not just limited to whom they educate but also with respect to schools' systemic factors.

Both public and private schools across the three countries vary in important ways in terms of school variables. Table 2 shows that the degree of reported competition and autonomy as well as staffing practices also differ by school type and could be additional pathways leading to the private-public learning differences. Private schools report facing more competition alongside greater levels of autonomy in decisions regarding staffing, budgeting, and instructional policies. For instance, school autonomy indices for instructional policies are between 0.25–0.75 larger in private institutions for the three countries.

Specifically, Australia shows overall high levels of competition and autonomy, and private-public differences are widest in Spain. In terms of principal-reported support and sanctions, private schools in Spain are especially different from their public counterparts. These distinctions are somewhat less pronounced in Portugal, followed by Australia. Teacher-reported support is generally high in Portugal, regardless of school type. Private schools often do better in terms of supporting their teachers. In terms of teacher-reported sanction, especially dismissal of under-performing teachers, private schools appear as more likely to exercise this form of sanction. Table 2, thus, concurs with some of the standard literature on the mechanisms and processes where public and private schools differ.

5.2. OLS estimates for the private school effect

Having outlined the distinctions by school type in terms of who they enroll and their internal mechanisms, in Table 3 we present OLS estimates on public and private performance differences. Controlling for background attributes, the left panel of this table presents the results for the full sample. Controlling for student and home background, the private school coefficient is still statistically significant for maths, reading, and science for the three countries. Effects are particularly pronounced in Australia and Portugal (around 18-30 and 24-32, respectively), even though students attending private schools in Spain achieve between 14-17 points above public-school students.

[Table 3 here]

We next present the results of the matched sample in the right panel of Table 3. Full details on the matching procedure can be found in Appendix A.[10] As expected, matching amounted to a reduction in sample size in each country. In Australia and Portugal, the coefficient on private is no longer statistically significant for maths, with the private school effect being of similar magnitude as for the full sample. Overall, the private school effect post-matching for each country remains statistically significant, regardless of the outcome measured.

5.3. OLS estimates of systemic school factors

Before presenting the key empirical analysis on the relationship of systemic school factors with learning gap across the learning distribution in the next section based on matched sub-samples, this section includes OLS impacts of systemic school factors. These estimates show the role of systemic school factor for the average performer. For completeness, we carry the analysis not only for the matched sub-samples, but also for the full sample (using the same students' and families' controls as in the previous section).

To begin with, Table 4 contains results for the full sample. Overall, across the five groups of systemic factors and the three countries, the private-public learning difference persists. The private school dummy only lacks statistical significance (at 10%) for maths in Australia (for competition, autonomy and school principal factors). In Spain, for the three outcomes for school principal factors while in Portugal the learning persists for all school systemic factors.

[Table 4 here]

In Table 5 we present the first analysis that motivates the paper, that is, whether systemic differences between public and private schools explain private-public learning gaps after isolating individual and household factors with a matching approach. Here, using the matched sub-samples derived by the matching approach, results for the private school dummy are presented controlling for each school systemic factor.

Estimates in Table 5 suggest that the greater competitive pressure faced by private schools may be important in explaining private-public performance gaps in science in Australia and Portugal, but not for reading. Greater school autonomy reported by private schools in Australia and Spain similarly appears to be important in explaining the private-public performance gaps in these systems. But this is not true for Portugal where gaps in private-public autonomy are somewhat smaller (as shown in Table 2). Principal-reported support and sanctions both appear more important in Spain, where there are considerably different levels of support, promotion, and dismissal policies in private schools (see Table 2). Teacher-reported sanction and support are less consistently related with explaining the private-public performance gap, except in Portugal, where the generally higher levels of support received by private school teachers seem to explain the learning performance gap.

[Table 5 here]

A comparison of the estimates of Tables 4 and 5 shows the empirical importance that isolating students' and households' factors with a matching approach has in our application.

In fact, qualitative different conclusions on the role of some systemic school factors are obtained with either analysis. For instance, estimates for positive private-public learning gap for competition (science-Australia, and maths- and science- Portugal) and teacher support (Portugal) are significant in the full sample (with controls) but they are not in the matched sub-sample analysis.

5.4. Counterfactual decomposition estimates of systemic school factors for matched sub-samples

This section contains the second analysis that motivates the paper. Here, by using a counterfactual decomposition of the learning gap based on matched sub-samples, we attempt to answer the following questions for comparable private and public students: Would private school students perform worse if they had the same distributions of systemic factors that public school's students? Does the importance of school systemic factors vary by students' performance level, and if so, what is behind this variation –either the effect of the systemic school factor per se or the aggregated level or contextual effect differences between private and public schools? To answer these questions, we run two sets of analyses. In the first analysis, we look at the total effect of school systemic factors as a whole, dividing the total impact between differences in coefficients (or the return to student's characteristics) and differences in the aggregate level of the school systemic factors (or inputs). In the second analysis, we look at the relationship at the level of each specific systemic school factor with the private-public learning gap.

Results for the first analysis −showing the decomposition of the private-public learning gap into changes in covariates and into changes in coefficients grouping all school systemic factors− are displayed in Figure 1 Panel A (maths) and Panel B (reading). The y-axis measures the performance that private students would have had if they had attended schools with the same distribution of systemic characteristics (and their associations with performance) than public school students, while the x-axis denotes the quantiles of the learning distribution. But due to the construction of the decomposition (see equation (3)), the interpretation of the counterfactual is in the opposite direction (sign). Hence, for a given quantile, a positive estimate implies that private school students would have reached a lower learning score under the learning profile of public school students, while a negative estimate implies that their achievement would have been higher (for characteristics).

For both maths and reading, we find that most of the total positive private learning gap we found in Section 5.2 can be explained by the effect attributed to the covariates (for Australia and Spain) and for the coefficients (Portugal). This implies that in Australia and Spain private students perform better mainly because of the systemic factors they encounter in private schools where as in Portugal their better performance may be largely due to their own characteristics. In addition, covariates' effects are increasing with performance for Australia and rather constant for Spain, while coefficient effects are decreasing with performance (for the three countries, though particularly for Australia and Portugal). These patterns are more pronounced for maths than for reading.

[Figure 1]

Based on 90% confidence intervals, covariates' effects are positive and significant for all students regardless of their performance (in Spain) and for the 50% top half of students based on their performance distribution (in Australia). This suggests that all students and those students with above average performance attending private schools in Spain and Australia, respectively, would have performed worse if they had attended schools with the same level of systemic school factors than public schools. Put differently, the joint effect of the level of competition, autonomy and staffing policy acts as a distinct feature within private schools leading to increasing performance for those with larger ability in Australia and for all students in Spain, but not in Portugal. At the other end of scores' distributions, estimates suggest that low performing students attending private schools would have been worse off if they had the same returns to characteristics (same coefficient) of the worst-performing students from public schools, especially in Australia and Portugal.

Because we find that the private learning gap is mainly linked to the aggregate value or contextual level of the whole set of systemic factors, it is then important to know which of the competition, autonomy and staffing factors is contributing more to this gap. This second part of the counterfactual analysis, focusing on the gap explained for the covariate of each school systemic factor, is included in Figure 2 (maths) and Figure 3 (reading), showing similar patterns for the two outcomes.

We find that, remarkably, competition has a minor (and constant) effect (below 10 points) only for two countries in (i.e., Portugal and Spain – Panel A), whereas autonomy leads to consistently large effects for Spain for the whole distribution and for the group of high performing students in Australia, and its impact is negative for Portugal (Panel B). The degree of support that either school principals or teachers receive (Panels C and F) does not seem to make a difference on the private-public learning gap, hovering around zero. Yet, for Spain, the covariate school principal support is associated with considerable larger achievement (around 20 points) for private school students, decreasing as we move towards upper quantiles. Accountability is associated to the positive gap for private school students in Australia (teacher sanctions – Panel E) and Spain (school principal sanction – Panel D); in either case being relatively more beneficial for better performing private school students.

6. Conclusions and discussions

Scholars have argued that private schools perform better (or will perform better) than public schools due to greater competitive pressures, greater autonomy and different staffing practices. In this paper we investigated these hypotheses. We used the unique TALIS-PISA link data for Australia, Portugal and Spain which allowed us to estimate private-public performance gap and to inspect whether differences in systemic school factors, namely, competitive pressures, autonomy, and staffing practices may help explain this performance gap. Given the widely acknowledged and observed gaps in the private and public students' home backgrounds we used matching approaches to identify a sub-sample of public and private students in each country who exhibited the same observable student and family characteristics. We also conducted counterfactual decomposition analysis in addition to OLS analysis to investigate if the effect of systemic school factors varies by whether a student is a low or a high performer to understand more closely the potential mechanisms through which private schools may (or may not) outperform public schools.

When looking at the effect that the level of specific systemic school factors has on the private-public learning gap we found that, across the three countries, the mechanisms that are traditionally forwarded to explain private school performance are not consistently significant in explaining their performance in the matched sub-samples –which is the most comparable sample of private's and public's school students. For instance, we found that competition has a small effect and additional support provided to teachers is not linked to differential performance by private's and public's school students. This, of course, may be indicative of inadequate measures of these institutional factors. But these findings equally support the idea that the mediating effect that certain school systemic factors have are more related to family inputs, with their effects disappearing after a robust analysis is carried out through matched sub-samples.

Importantly, these findings point to the difficulty of identifying mechanisms that schools may ultimately adopt to change their performance. In this regard these findings are aligned with an extensive body of literature on school effects which highlights the difficulty of identifying school factors that are important in explaining variation in student performance (see Chudgar and Luschei, 2009, for a review). Our findings which focus more on school systemic factors or *processes* rather than *traditional inputs* provide additional nuances to this literature (see also Norton Grubb, 2008 for a discussion of types of school resources).

However, to the extent that school systemic factors are important, we find that a differentiation at which level accountability is measured is important to gain further insights into what drives the private-public learning gap. We found that teacher sanction is a consistent factor in explaining the private-public learning gap in Australia only, principal's sanction is relevant only in Spain, while in Portugal accountability measured in either form is negatively correlated to the learning gap. In country-cases such as Australia, where the counterfactual decomposition pointed to larger gradients by students' performance, accountability at the classroom-teacher level matters more, benefitting high performing students. In cases where learning is more homogenous, as we found in Spain, accountability at the school level (i.e., school principals) is more important in explaining the learning gap. The implications of these findings deserve closer attention.

The varying importance of school systemic factors to high- and low-performing student is one of the key contributions of our work. Although some systemic factors might be individually as powerful determinants of the private-public learning gap, their joint effect appears to lead to substantial gains for the population of high ability students. This result underscore the importance of paying attention to the learning levels of students attending private and public schools to isolate the importance of school systemic factors Their impact, when present, appears to depend on how ability-segregated an education system is.

Hence, our analysis points to a vital yet unexplored dimension in the school choice literature in the cross-national context: private school effects are heterogeneous across the learning distribution with the varying importance of school systemic factors for low, medium and high achieving students. We found that, competition, accountability and staffing policies play a limited role in narrowing the private-public gap among low performers but the reverse is true for high performers. The paper's results highlight an overreaching message: private school systems and processes studied here (competition, accountability and staffing policies) may be important to explain variation in private-public performance but perhaps only for the high performing students. The same "private school practices" when adopted for low performing students do not appear to yield greater learning levels. These findings have vital equity implications especially as policymakers consider growth in the private school sector and/or adopting similar practices of competition, accountability in the public-school sector.

Standard caveats apply to our conclusions. First, our sample consists of three countries and hence we cannot claim that all conclusions will hold for a different group of countries. Indeed, we highlight the country-dependency nature of our results and the need for carrying out further research not only with a different group of countries, but also with a different set of institutional factors. Second, even though our estimates are based on robust techniques, we cannot claim our results denote causality; instead, they show conditional associations.

References

Albrecht, J., van Vuuren, A., & Vroman, S. (2009). Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands. *Labour Economics*, 16(4), 383–396.

Altonji, J., T. Elder, and C. Taber. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113 (1), 151–184.

Austin, B., O. Adesope, B. French, C. Gotch, J. Bélanger, and K. Kubacka. (2015). Examining School Context and Its Influence on Teachers: Linking Talis 2013 with PISA 2012 Student Data. *OECD Education Working Papers* No. 115, OECD Publishing, Paris.

Battistin, E., & Meroni, E. C. (2016). Should we increase instruction time in low achieving schools? Evidence from Southern Italy. *Economics of Education Review*, *55*, 39-56.

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does management matter in schools?. *The Economic Journal*, *125*(584), 647-674.

Chen, M., V. Chernozhukov, I. Fernández-Val, and B. Melly, B. (2016). Counterfactual: An R Package for Counterfactual Analysis. arXiv:1610.07894.

Chernozhukov, V., I. Fernández-Val and B. Melly (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.

Chudgar, A., and Luschei, T. (2009). National Income, Income Inequality, and the Importance of Schools: A Hierarchical Cross-National Comparison, *American Educational Research Journal,* 46 (3), 626-658

Chudgar, A., and E. Quin. (2012). Relationship Between Private Schooling and Achievement: Results From Rural and Urban India. *Economics of Education Review* 31 (4), 376–390.

Chubb, J. E., and T. M. Moe. (1990). America’s Public Schools: Choice Is a Panacea. *The Brookings Review* 8 (3), 4–12.

Dehejia, R., and S. Wahba. (1999). Causal Effects in Nonexperimental Studies: Reevaluation of the Evaluation of Training Programs. *Journal of the American Statistical Association* 94, 1043–1062.

Delprato, M., Akyeampong, K., and Dunne, M. (2017). The impact of bullying on students' learning in Latin America: a matching approach for 15 countries, *International Journal of Educational Development*, 52, 37–57.

Dronkers, J., & Robert, P. (2008). Differences in scholastic achievement ofpublic, private government-dependent, and private independent schools. *Educational Policy*, 22(4), 541–577.

Evans, W.N. & Schwab, R.M. (1995). Finishing high school and starting college: do Catholic schools make a difference? *Quarterly Journal of Economics*, 110(4): 941-974.

Ferrera, J. M. C., and Izquierdo, M. G. (2016). TALIS-PISA link: Guidelines for a robust quantitative analysis. In *International Conference on Qualitative and Quantitative Economics Research (QQE). Proceedings* (p. 34). Global Science and Technology Forum.

Firpo, S., Fortin, N., and T. Lemieux, (2007). Unconditional Quantile Regressions, *Econometrica*, 77, 953–973.

Fortin, N., T. Lemieux, and S. Firpo (2011). Decomposition Methods in Economics, Volume 4 of Handbook of Labor Economics, Chapter 1, pp. 1-102. Elsevier.

Freeman, C., O'Malley, K., & Eveleigh, F. (2014). Australian teachers and the learning environment: An analysis of teacher response to TALIS 2013.

Friedman, M. (1955). *The Role of Government in Education*. New Brunswick, NJ: Rutgers University Press.

Frölich, M., and B. Melly (2010). Estimation of quantile treatment effects with Stata, *The Stata Journal*,10(3), 423–457.

Frölich, M., and B. Melly (2013). Unconditional quantile treatment effects under endogeneity, *Journal of Business & Economic Statistics*, 31 (3), 346-357.

Grubb, W. N. (2008). Multiple resources, multiple outcomes: Testing the "improved" school finance with NELS88. *American Educational Research Journal*, *45*(1), 104-144.

Gumus, S. (2013). The effects of teacher-and school-level factors on teachers' participation in professional development activities: The role of principal leadership. *Journal of International Education Research*, *9*(4), 371.

Guo, S., and M. Fraser. (2010). *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: Sage Publications.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24(3),1141–1177.

Hanushek, E.A., Woesmann, L., (2008). Education Quality and Economic Growth. World Bank, Washington, DC.

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212-232.

Hooge, E., T. Burns and H. Wilkoszewski (2012), "Looking Beyond the Numbers: Stakeholders and Multiple School Accountability", *OECD Education Working Papers*, No. 85, OECD Publishing.

Hoxby, C. M. (2001). How School Choice Affects the Achievement of Public School Students. *Choice With Equity* 150.

Hoxby, C. M. (2002). Would School Choice Change the Teaching Profession? *The Journal of Human Resources* 37 (4), 846–891.

Iacus, S., King, G., and Porro. G. (2011). Causal Inference Without Balance Checking: Coarsened Exact Matching. Political Analysis http://j.mp/iUUwyH

Kaplan, D., & McCarty, A. T. (2013). Data fusion with international large scale assessments: a case study using the OECD PISA and TALIS surveys. *Large-scale Assessments in Education*, *1*(1), 6.

Le, A.T. and Miller, P. (2003), 'Choice of school in Australia: Determinants and consequences', *Australian Economic Review*, 36: 55 – 78.

Leuven, E., and B. Sianesi. (2012). PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Statistical Software Components. Boston College Department of Economics.

Leithwood, K. A., and C. Riehl. (2003). *What We Know About Successful School Leadership*. Nottingham: National College for School Leadership.

Lounkaew, K. (2013). Explaining urban–rural differences in educational achievement in Thailand: Evidence from PISA literacy data. *Economics of Education Review* 37, 213-225.

Lubienski, C. (2004). Charter School Innovation in Theory and Practice: Autonomy, R&D, and Curricular Conformity. Taking Account of Charter Schools: What's Happened and What's Next 72–90.

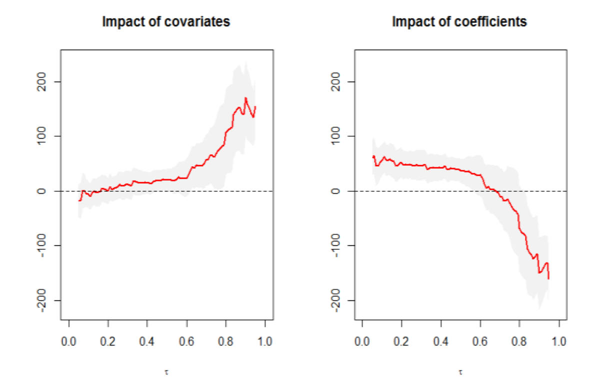
Machado, J. A. F., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. Journal of Applied Econometrics, 20(4), 445–465.

Melly, B. (2006). Estimation of counterfactual distributions using quantile regressions. Working Paper. University of St. Gallen.
Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. International economic review, 693–709.
OECD (2011). PISA in focus. 7 (August).
OECD (2012a). *PISA 2012 Technical Report*. Paris, France.
OECD (2012b), Public and Private Schools: How Management and Funding Relate to Their Socio-economic Profile, OECD Publishing. http://dx.doi.org/10.1787/9789264175006-en
OECD (2013a). Teaching and Learning International Survey (TALIS) 2013 Technical Report. Paris.
OECD (2013b). TALIS 2013 User Guide to the International Database. Paris.
OECD (2013c). Education Policy Outlook: Australia, Paris.
OECD (2014a). Education Policy Outlook: Spain, Paris.
OECD (2014b). Education Policy Outlook: Portugal, Paris.
Oliveira, P. R., Belluzzo, W. & Pazello, E. T., (2013). The private-public test score gap in Brazil. Economics of Education Review, Elsevier, 35(C), 120-133.
Ozga, J. (2013). Accountability as a policy technology: accounting for education performance in Europe. *International Review of Administrative Sciences*, *79*(2), 292-309.
Patrinos, H. (2013). Private education provision and public finance: the Netherlands, *Education Economics* 21(4), 392–414.
Podgursky, M. (2006). Teams Versus Bureaucracies: Personnel Policy, Wage-setting, and Teacher Quality in Traditional Public, Charter, and Private Schools. Education Working Paper Archive.
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458.
Robinson, J., and Lubienski, S. (2011). The development of gender achievement gaps in math ematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings, *American Educational Research Journal* 48(2), 268-302.
Rosenbaum, P., and D. Rubin. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, 41–45.
Rubin, D. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology* 2, 169–188.
Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, *11*(1), 45-61.
Sealy, K. M., Perry, S. M., and DeNicola, T. C. (2016). Relationships and Predictors of Principal Job Satisfaction Across Multiple Countries: A Study Using Talis 2013 and PISA 2012.
Vandenberghe, V., and Robin, S. (2004). Evaluating the effectiveness of private education across countries: A comparison of methods. *Labour Economics*, 11, 487-506.
Vieluf, S., Kunter, M., & van de Vijver, F. J. (2013). Teacher self-efficacy in cross-national perspective. *Teaching and Teacher Education*, *35*, 92-103.
West, M.R. and Woessmann, L. (2010), Every Catholic child in a Catholic school': Historical resistance to state schooling, contemporary private competition and student achievement across countries, *Economic Journal*, 120 (546), F229–F255
Woessmann, L. (2007). International evidence on school competition, autonomy, and accountability: A review. *Peabody journal of education*, *82*(2-3), 473-497.

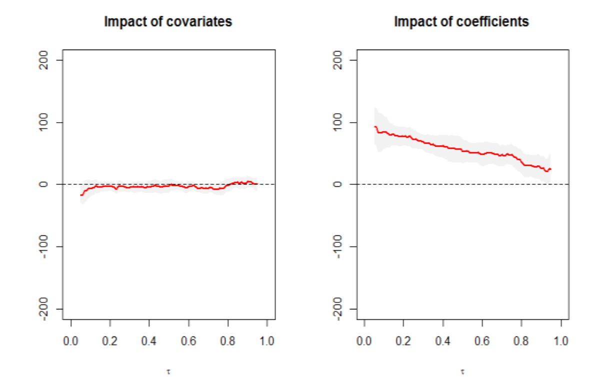
Figure 1. Impact of covariates and coefficients of systemic school factors on the private-public learning gap based on matched sub-sample.

Panel A – Maths scores

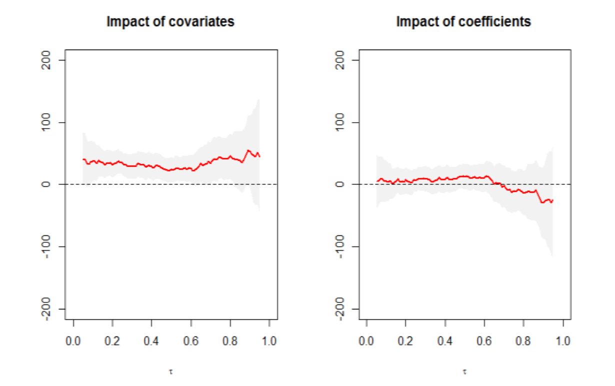
Australia



Portugal

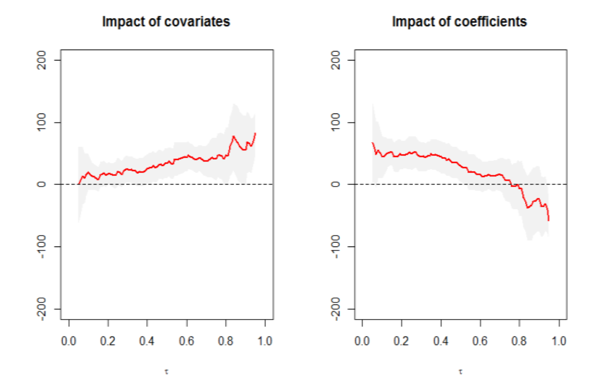


Spain

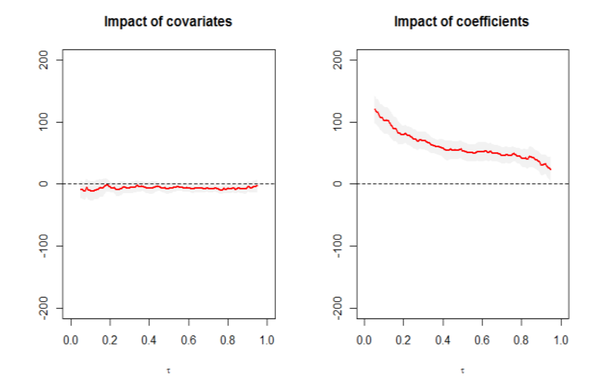


Panel B – Reading scores

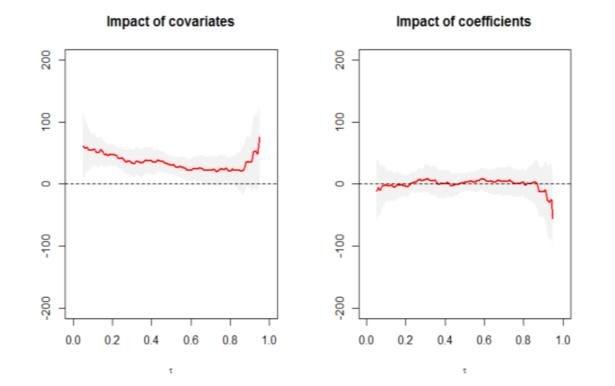
Australia



Portugal



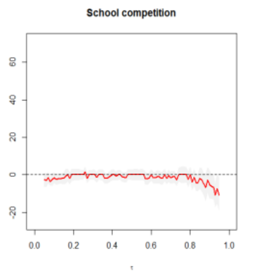
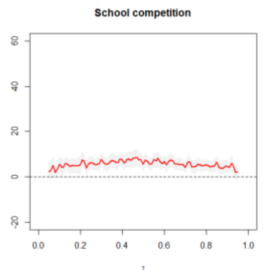
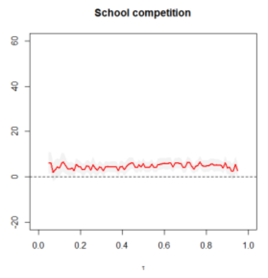
Spain



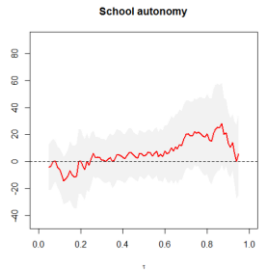
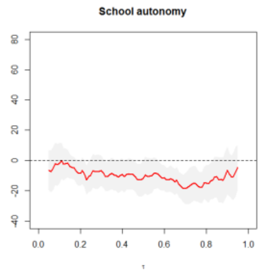
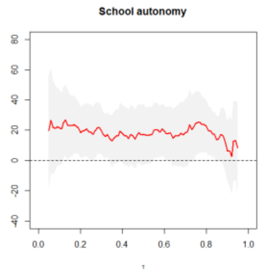
Notes: (1) Counterfactual decomposition of total private school effect by the effect of covariates and coefficients using logistic distribution regression where the counterfactual population is defined by the private school dummy. (2) Weighted bootstrap (R=200) is used for the construction of 90% uniform confidence intervals. (3) Estimates are based on the R package counterfactual (Chen et al., 2016). (4) The y-axis denotes the learning profile of private students under the hypothetical case of having the same distribution of systemic characteristics of public school students. The x-axis denotes quantiles of the learning score distribution. (5) Interpretation. The total difference of the private-public learning gap is decomposed as the sum of changes into covariates (systemic factors) and changes into coefficients. First sub-plot: negative (positive) values for the covariates implies that private school's students would perform better (worse) if they had the same distribution the public schools' students do have for the covariates. Second sub-plot: the same interpretation but with respect to the relationship (coefficient) of systemic factor with the private-public learning gap. (6) Illustrative example – Australia (Panel A – Maths scores). In the first sub-plot for covariates we can see that there is a significant effect (where the confidence interval does not overlap with zero) in the top 40% of the distribution (for $\tau \geq 0.60$). This implies that only higher performers from private schools would have been worse-off if they had the same distribution of systemic factors than public school students. In the second plot, for coefficients, we can see that lower performers (in the bottom 60% of the distribution) from private schools would be worse-off if they had the same associations of characteristics with performance than public school students.

Figure 2. Impact of specific covariates of systemic school factors on the private-public learning gap – Maths scores

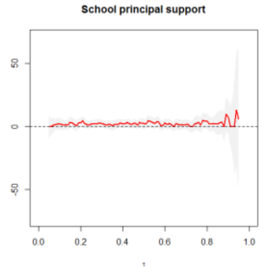
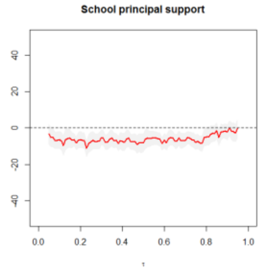
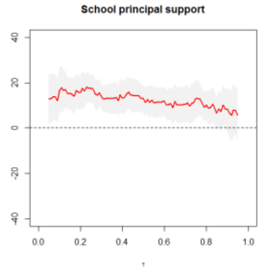
Panel A – School competition

Australia	Portugal	Spain
		

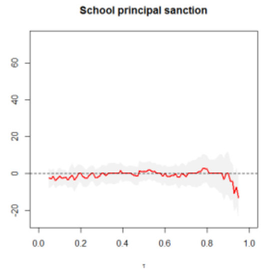
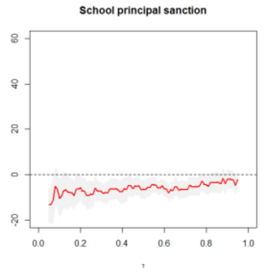
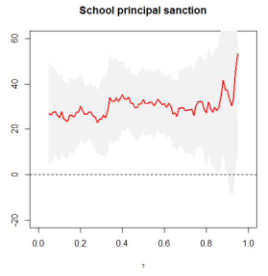
Panel B – School autonomy

Australia	Portugal	Spain
		

Panel C – School principal support

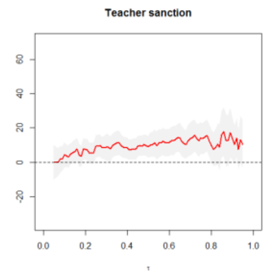
Australia	Portugal	Spain
		

Panel D – School principal sanction

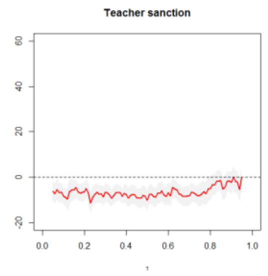
Australia	Portugal	Spain
		

Panel E – Teacher sanction

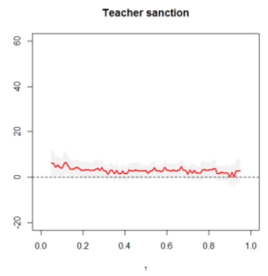
Australia



Portugal

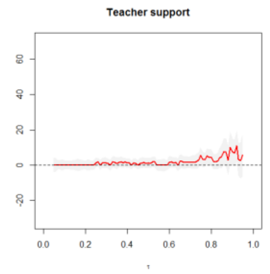


Spain

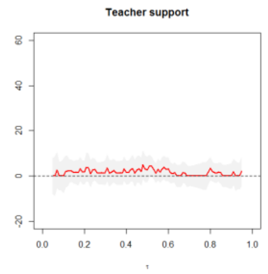


Panel F – Teacher support

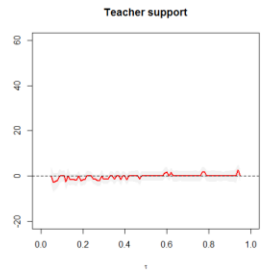
Australia



Portugal



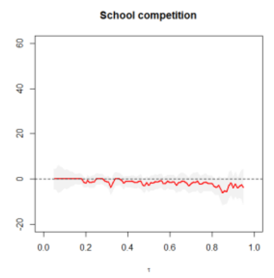
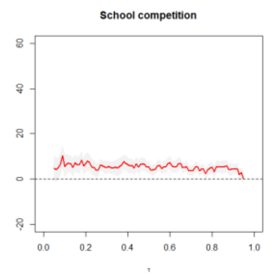
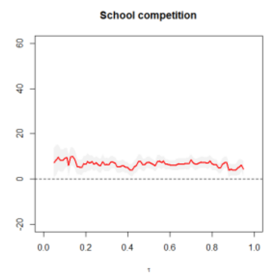
Spain



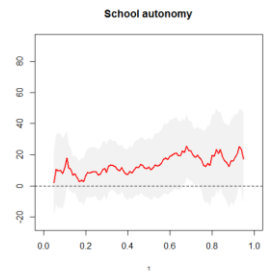
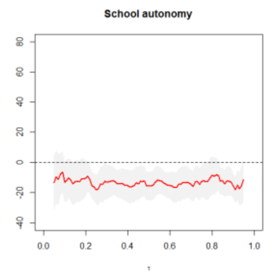
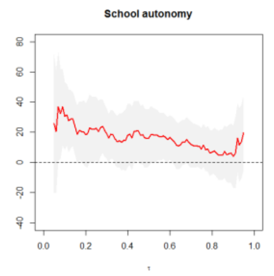
Notes: (1) See notes in Figure 1.

Figure 3. Impact of specific covariates of systemic school factors on the private-public learning gap – Reading scores.

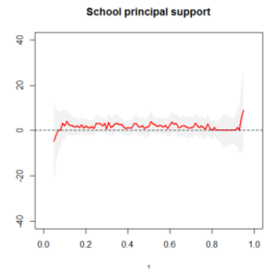
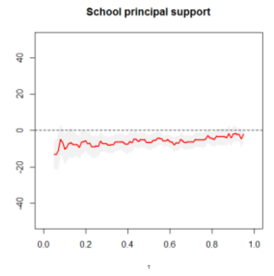
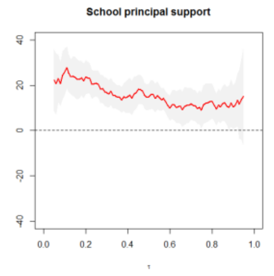
Panel A – School competition

Australia	Portugal	Spain
		

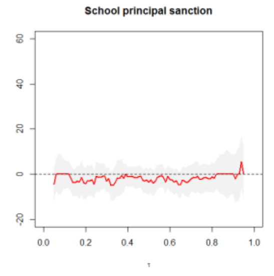
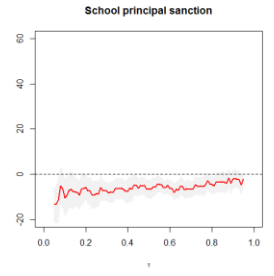
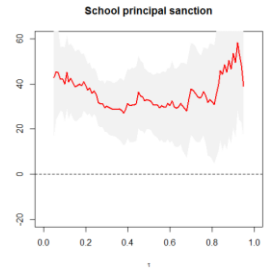
Panel B – School autonomy

Australia	Portugal	Spain
		

Panel C – School principal support

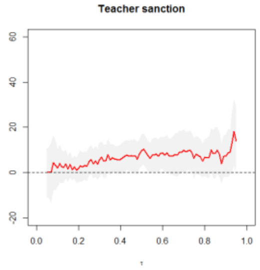
Australia	Portugal	Spain
		

Panel D – School principal sanction

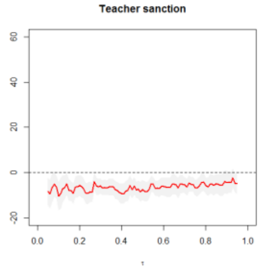
Australia	Portugal	Spain
		

Panel E – Teacher sanction

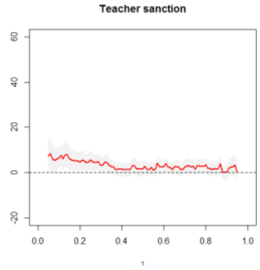
Australia



Portugal

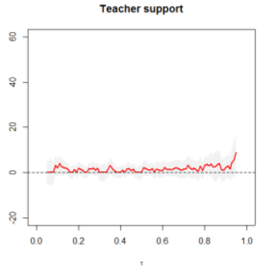


Spain

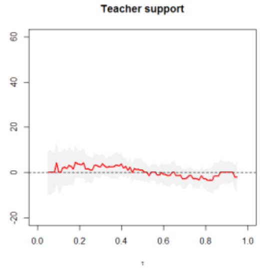


Panel F – Teacher support

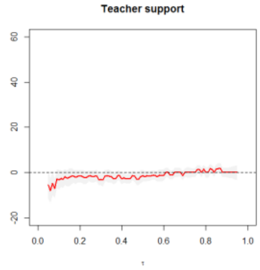
Australia



Portugal



Spain



Notes: (1) See notes in Figure 1.

Tables

Table 1. *Summary statistics for tests scores and explanatory variables*

	Australia		Portugal		Spain	
	Public	Private	Public	Private	Public	Private
Test scores outcomes						
Maths	494.42	527.86	482.09	526.22	470.08	509.43
Reading	497.14	543.26	483.44	529.61	476.36	510.93
Science	506.07	548.76	483.64	527.00	484.68	519.15
Covariates: students						
Age	15.84	15.84	15.79	15.83	15.90	15.90
Male	0.53	0.49	0.50	0.55	0.50	0.53
Language of test	0.88	0.94	0.97	0.99	0.77	0.79
Mother work	0.72	0.78	0.73	0.74	0.65	0.74
Wealth	0.38	0.80	0.18	0.56	-0.05	0.15
Urban	0.84	0.93	0.59	0.34	0.69	0.93
Mother education level	3.29	3.47	2.27	2.66	2.75	3.21
Father education level	3.09	3.32	2.08	2.49	2.64	3.04
Number of books	3.24	3.61	2.70	3.01	3.15	3.59
Truancy	1.52	1.40	1.73	1.61	1.51	1.40
Nuclear family	0.84	0.89	0.89	0.91	0.90	0.90
Child immigrant	0.89	0.93	0.93	0.96	0.88	0.93
Number of observations	1,382	1,030	3,772	307	5,832	3,038
Number of schools	68	54	128	12	202	108

Notes: (1) Student weighted means based on PISA 2012. (2) Non-cognitive outcomes are scaled continuous indices derived in the dataset (see OECD [2012a] for details for items for each index). (3) Wealth is a continuous index based on students' responses on whether they had various assets at home: own room, internet, televisions, computers, cars, etc. (4) Parental education is a discrete variable based on highest schooling: 1 (ISCED 1 or less), 2 (ISCED 2), 3 (ISCED 3B, 3C) and 4 (ISCED 3A). (5) Truancy measures the frequency of being late for school: 1 (none), 2 (one or two times), 3 (three or four times) and 4 (five or more times). (6) Nuclear family measures a two-parent family.

Table 2. *Summary statistics of school and teacher factors for matched sub-sample analysis*

	Australia		Portugal		Spain	
	Public	Private	Public	Private	Public	Private
School competition	1.76	1.99	1.34	1.89	1.23	1.81
School autonomy						
For staffing	2.13	2.82	2.14	2.51	1.09	2.66
For budgeting	1.08	1.86	1.09	1.54	1.03	1.42
For instructional policies	2.37	2.81	1.84	2.09	1.46	2.21
School principal - support						
Measures to remedy weaknesses in teaching	2.16	2.30	2.06	2.42	1.99	2.21
Development or training plan for each teacher	2.23	2.18	1.57	1.97	1.35	1.99
Mentor to help the teacher	2.20	2.07	1.44	1.73	1.24	1.79
Change in a teacher’s work responsibilities	1.61	1.73	1.35	1.37	1.34	1.77
Teacher - support						
Training plan is established to improve teacher's work	0.61	0.45	0.42	0.76	0.46	0.60
Feedback is provided on a thorough assessment of teaching	0.29	0.34	0.65	0.84	0.23	0.27
Measures to remedy weaknesses in teaching are discussed	0.62	0.57	0.62	0.75	0.57	0.72
Mentor to help the teacher	0.48	0.44	0.51	0.40	0.17	0.22
School principal - sanction						
Material sanctions for poor performing teacher	1.12	1.06	1.06	1.00	1.03	1.02
Change teacher's salary or financial bonus	1.05	1.19	1.02	1.24	1.04	1.12
Change in teacher’s promotions	1.70	1.83	1.30	1.55	1.13	1.64
Dismissal or non-renewal of contract	1.64	1.80	1.26	1.63	1.07	1.73
Teacher - sanction						
The best performing teachers receive recognition	0.28	0.21	0.22	0.11	0.15	0.28
Appraisal and feedback have impact on teaching	0.55	0.50	0.57	0.85	0.39	0.53
Under-performing teachers are dismissed	0.12	0.43	0.38	0.60	0.11	0.33
Number of observations	1,382	1,030	3,772	307	5,832	3,038
Number of schools	68	54	128	12	202	108

Notes: (1) School competition is a discrete variable measuring the degree of competition 0 (none), 1 (one school), 2 (2 or more). (2) School autonomy comprises three indices measuring autonomy in staffing, budgeting, and instructional policies. (3) For further details on the remaining variables see OECD (2013a).

Table 3. *OLS estimates for private school dummy in full and matched samples*

	Full sample			Matched sample		
Outcome	Australia	Portugal	Spain	Australia	Portugal	Spain
Maths	18.07**	23.86***	17.00***	14.57	20.53	18.57***
Reading	30.48***	31.73***	14.18**	29.03***	29.94**	15.16**
Science	28.15***	26.07***	14.09***	26.33**	24.79*	14.62***
N	1,954	3,558	7,613	1,396	566	4,986

Notes: (1) All estimates include the students' controls and school location of Table 1. (2) Further details on the matching procedure and its' quality can be found in Appendix A. (3) School clustered standard errors.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. *OLS estimates for private school dummy with additional systemic school factors. Full sample.*

	Australia			Portugal			Spain		
	Maths	Reading	Science	Maths	Reading	Science	Maths	Reading	Science
Panel A: Competition									
Private - OLS	14.18	27.54***	24.36**	20.28***	29.69***	26.71***	16.02***	12.59**	12.13**
N	1,565	1,565	1,565	3,053	3,053	3,053	7,455	7,455	7,455
Panel B: Autonomy									
Private - OLS	8.32	20.90**	15.39	26.90***	37.15***	29.89***	13.39*	29.40**	16.55
N	1,593	1,593	1,593	3,309	3,309	3,309	7,215	7,215	7,215
Panel C: School principal - support									
Private - OLS	15.07	28.02***	24.64***	26.43***	31.26***	29.68***	3.10	-3.73	0.48
N	1,579	1,579	1,579	3,377	3,377	3,377	4,821	4,821	4,821
Panel D: School principal -sanction									
Private - OLS	14.03	29.91***	26.71***	28.71***	33.42***	29.85***	-1.12	-3.30	-2.57
N	1,579	1,579	1,579	3,381	3,381	3,381	4,821	4,821	4,821
Panel E: Teacher-sanction									
Private - OLS	17.73*	28.93***	24.91***	35.09***	44.68***	32.84***	15.83***	14.09**	12.35**
N	1,759	1,759	1,759	3,287	3,287	3,287	7,074	7,074	7,074
Panel F: Teacher-support									
Private - OLS	15.76*	29.20***	22.92***	22.06***	30.66***	22.37***	17.11***	16.12***	13.39***
N	1,772	1,772	1,772	3,280	3,280	3,280	7,019	7,019	7,019

Notes: (1) All estimates include the students' controls and school location of Table 1 plus the specific school systemic factor. (2) School clustered standard errors.
Significance levels: * $p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

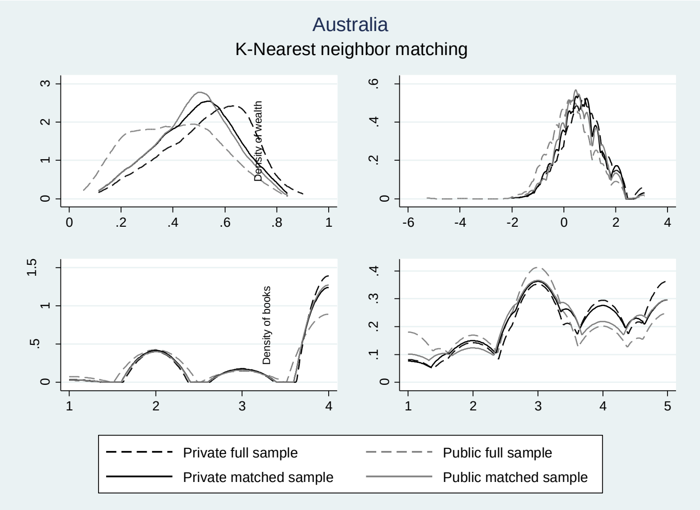
Table 5. Explaining the private gap on learning scores. School systemic factors. OLS estimates based on matched sub-samples

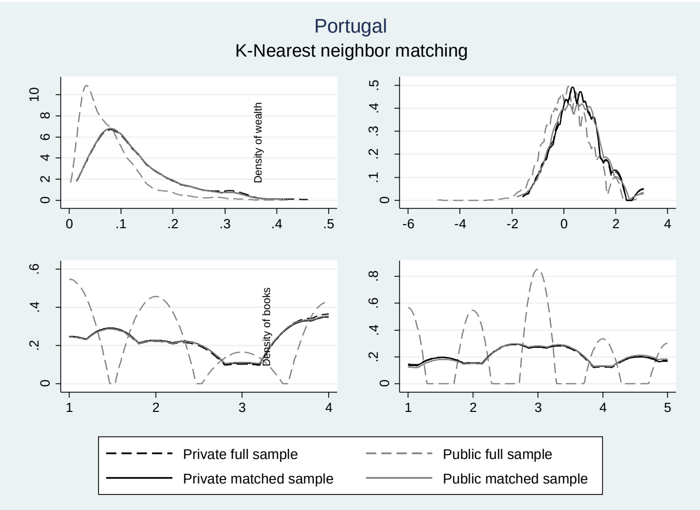
	Australia			Portugal			Spain		
	Maths	Reading	Science	Maths	Reading	Science	Maths	Reading	Science
Panel A: School competition									
Private - OLS	8.51	25.83**	21.01	12.21	22.43**	19.57	18.26***	15.10**	13.95**
N	1,108	1,108	1,108	503	503	503	4,871	4,871	4,871
Panel B: School autonomy									
Private - OLS	-2.37	14.70	9.78	32.73**	41.91***	33.89**	7.25	24.10*	8.69
N	1,136	1,136	1,136	550	550	550	4,660	4,660	4,660
Panel C: School principal - support									
Private - OLS	10.09	27.40**	22.21*	24.15*	30.77**	33.94**	-2.04	-6.30	-5.85
N	1,114	1,114	1,114	528	528	528	3,327	3,327	3,327
Panel D: School principal -sanction									
Private - OLS	9.24	28.70**	23.86*	22.15	28.19*	29.14	-4.06	-8.00	-6.68
N	1,114	1,114	1,114	526	526	526	3,327	3,327	3,327
Panel E: Teacher-sanction									
Private - OLS	17.04	30.07***	25.19**	18.01	30.87**	17.39	16.72***	14.57**	11.90**
N	1,265	1,265	1,265	461	461	461	4,655	4,655	4,655
Panel F: Teacher-support									
Private - OLS	14.72	28.54***	22.13*	7.66	16.84	7.14	17.86***	17.07***	13.23**
N	1,271	1,271	1,271	493	493	493	4,628	4,628	4,628

Notes: (1) Each panel is a different regression where each group of covariates are introduced separately. (2) School clustered standard errors.
Significance levels: * p < 0.10, **p < 0.05, ***p < 0.01.

Appendix A: Quality of matching

Figure A1. *Distribution of covariates in full and matched sample by school type*





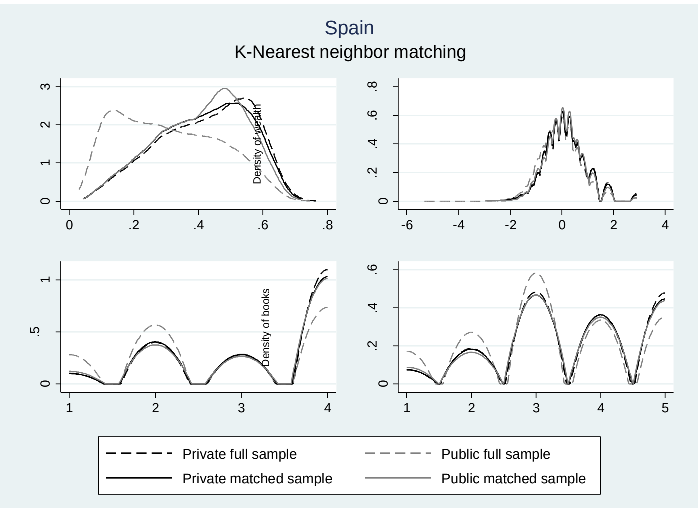


Table A1. Differences in means for matched sample for dependent variables, propensity score and covariates by public/private schools

	Australia			Portugal			Spain		
	Private	Public	t-test	Private	Public	t-test	Private	Public	t-test
Test scores outcomes									
Maths	517.34	496.19	4.18	534.38	504.59	4.31	514.83	500.10	6.47
Reading	533.30	498.74	7.10	537.87	498.14	6.02	518.88	503.83	6.49
Science	539.47	509.15	5.90	535.61	502.68	5.05	524.30	515.10	4.21
Propensity score	0.49	0.48	1.07	0.12	0.12	0.02	0.41	0.41	1.30
Covariates: students									
Age	15.85	15.84	0.37	15.86	15.87	-0.61	15.91	15.91	-0.15
Male	0.51	0.53	-0.59	0.54	0.55	-0.34	0.52	0.53	-0.68
Language of test	0.93	0.93	0.21	0.99	0.99	0.45	0.85	0.86	-0.99
Mother work	0.77	0.77	0.00	0.72	0.72	0.00	0.71	0.71	0.06
Wealth	0.64	0.59	1.26	0.55	0.52	0.40	0.14	0.11	1.43
Urban	0.90	0.89	0.44	0.37	0.39	-0.35	0.84	0.83	0.65
Mother education level	3.40	3.43	-0.60	2.60	2.61	-0.14	3.23	3.22	0.39
Father education level	3.16	3.21	-0.98	2.47	2.56	-0.87	3.09	3.09	-0.19
Number of books	3.49	3.44	0.75	3.04	3.12	-0.70	3.60	3.58	0.58
Truancy	1.49	1.48	0.40	1.66	1.72	-1.00	1.38	1.39	-0.56
Nuclear family	0.88	0.89	-0.50	0.92	0.93	-0.47	0.91	0.91	-0.79
Child immigrant	0.92	0.92	0.20	0.97	0.95	0.87	0.92	0.92	0.16
N	698	698		283	283		2493	2493	

Notes: (1) T-test values for significance differences between private and publics variables at 10% level of significance is less than 1.64, and for significance level at 5% less than 1.96 (absolute values). (2) Matched samples are based on nearest neighbour without replacement.

Table A2. Matching alternative approaches estimates and performance

	Australia				Portugal				Spain			
	Nearest neighbour		Radius	Kernel	Nearest neighbour		Radius	Kernel	Nearest neighbour		Radius	Kernel
	No rep	With rep			No rep	With rep			No rep	With rep		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Test scores estimates												
Maths	21.14***	14.94**	21.79***	21.11***	29.78***	27.94***	34.90***	35.95***	14.74***	17.87***	14.49***	15.02***
Reading	34.56***	26.13***	34.65***	34.25***	39.73***	39.87***	42.94***	42.89***	15.05***	18.67***	15.08***	15.29***
Science	30.33***	21.93***	31.09***	30.38***	32.94***	31.36***	36.74***	37.27***	9.20***	12.01***	9.01***	9.50***
Caliper/bandwidth	0.018	0.018	0.018	0.06	0.006	0.006	0.006	0.06	0.016	0.016	0.016	0.06
t-test (diff sign)	None	None	None	None	None	None	None	None	None	Two	None	None
Variance ratio	None	None	None	None	None	None	None	None	One	One	None	None
Mean bias (%)	2.7	3.8	1.3	1.2	4.1	4.8	0.9	3.8	1.5	2.7	0.7	1.0
B Rubin and R (%)	11.9, 0.97	16.1, 1.14	6.1, 0.97	5.9, 1.09	18.1, 0.90	21.1, 0.77	3.8, 1.27	15.2, 0.84	7.2, 1.06	12.0, 1.12	3.2, 1.02	4.0, 1.09

Notes: (1) The variance ratio (for continuous covariates) of treated over non-treated should equal 1 if there is perfect balance. None indicates there is perfect balance for all covariates, and one and two denote the number of covariates without perfect balance. (2) Mean bias is a summary indicator of the distribution of the absolute bias for all covariates, which ideally should be less than 5%. (3) Rubin's B and R statistics are absolute standarised differences on the mean and ratio of the propensity score in the treated and non-treated groups (see Rubin [2001] for details).
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Estimates accounting for unobservables

Another challenge for the identification of the private school effect is the non-random selection into private schools. Student's background characteristics that influence schooling decisions are likely to influence learning outcomes independently of school choice, since they would be related to other parental inputs (Altonji et al., 2005; West and Woessmann 2010). There is also the mechanism of student sorting into high-performing schools which will induce students to perform better because they will be in better schools (Patrinos, 2013). Any observed final educational gap can be due inherent differences in characteristics of students and their families. Some of these differences can be controlled for as they are observed (e.g., parental education) but others are unobserved (such as student's ability, motivation or parents' value of education and involvement).

Here, in Appendix B, to allow for the possibility of students and household unobservables driving school choice we employ an instrumental variable (IV) approach. Studies tend to use as an instrument (i.e., a phenomenon that is related to school choice but which does not have an impact on outcomes) family religion affiliation and to a lesser extent location or prevalence of specific schools (e.g., Evans and Schwab, 1995; Le and Miller, 2003). We employ two instruments: religion and religion interacted with residence restrictions on admission. The former instrument is already tested for the context of the private-public learning gap in Netherlands based on PISA (Patrinos, 2013), while the latter is a new empirical contribution which captures the fact that religious parents may choose to live in places near a religious school to increase the chances that their children would secure a place in them. Since it is quite possible for unobservables to have a different effect for low or high achievers, we use an IV approach but within a quantile treatment effects (QTEs) framework.

Based on a linear model, estimates for instruments' validity and tests seem to support our modelling approach overall –in particular, for the religion instrument z_1 (see Table B1). First-stage results indicate a statistically significant association of z_1 with private school attendance (i.e., parents who endorse the school's religious policy are more likely to send their children to private school), but not with outcomes (i.e., parental endorsement of the religious policy does not have a direct impact on scores) and test results show that instruments are not weak, though we find evidence of endogeneity only for Australia.

[Table B1 here]

Results for IV QTEs for the private school effect are displayed in Table B2. The focus here is on IV QTEs estimates of Panel B (using the religion instrument z_1). We also include the exogenous unconditional quantile estimates in Panel A for comparison. A few findings emerge from this table. First, the similarity of the full and sub-sample point estimates points towards the fact that unobservables are still at play after restricting the analysis to the subsample of public and private students with similar observed covariates. Second −and more importantly− selection bias operates differently between countries. Results for Australia suggest that the association learning-private school is driven by intrinsic students' and family's characteristics not included in our model (i.e., estimates are not significant for all quantiles for maths and above Q75 for reading), whereas for Portugal and Spain endogeneity only has in impact for the larger ability (in top quartile) group.

[Table B2 here]

Variation among countries' results is clearly shown in Figure B1 which, in turn, is supported by the fact we only found evidence of endogeneity for Australia possibly explained by low ability students self-selecting into the private school system and thereby increasing their achievement. In the other two countries, however, student's sorting into high performing schools seems a more plausible explanation since endogeneity only impacts at the top end of the achievement distribution.

[Figure B1 here]

Table B1. *IV (linear) first stage results and tests (full sample)*

	Maths		Reading	
	z1	z2	z1	z2
Australia				
First stage - effect of z on private school dummy	0.66***	0.63***	0.66***	0.63***
Endogeneity test	8.81	0.01	9.94	0.29
p-value	0.00	0.93	0.00	0.59
Weak identification	104.22	90.43	104.22	90.43
p-value	0.00	0.00	0.00	0.00
SWr2	0.44	0.35	0.44	0.35
Effect of z on y	-4.03	12.36	5.13	23.07***
Portugal				
First stage - effect of z on private school dummy	0.13**	0.28	0.13**	0.28
Endogeneity test	0.47	2.35	0.96	2.43
p-value	0.49	0.13	0.33	0.12
Weak identification	5.16	1.89	5.16	1.89
p-value	0.03	0.17	0.03	0.17
SWr2	0.05	0.03	0.05	0.03
Effect of z on y	8.80	-15.78	12.31	-14.24
Spain				
First stage - effect of z on private school dummy	0.64***	0.62***	0.64***	0.62***
Endogeneity test	0.05	0.24	0.29	1.29
p-value	0.82	0.62	0.59	0.26
Weak identification	179.71	82.35	179.71	82.35
p-value	0.00	0.00	0.00	0.00
SWr2	0.28	0.07	0.28	0.07
Effect of z on y	9.93**	3.35	7.34	-4.38

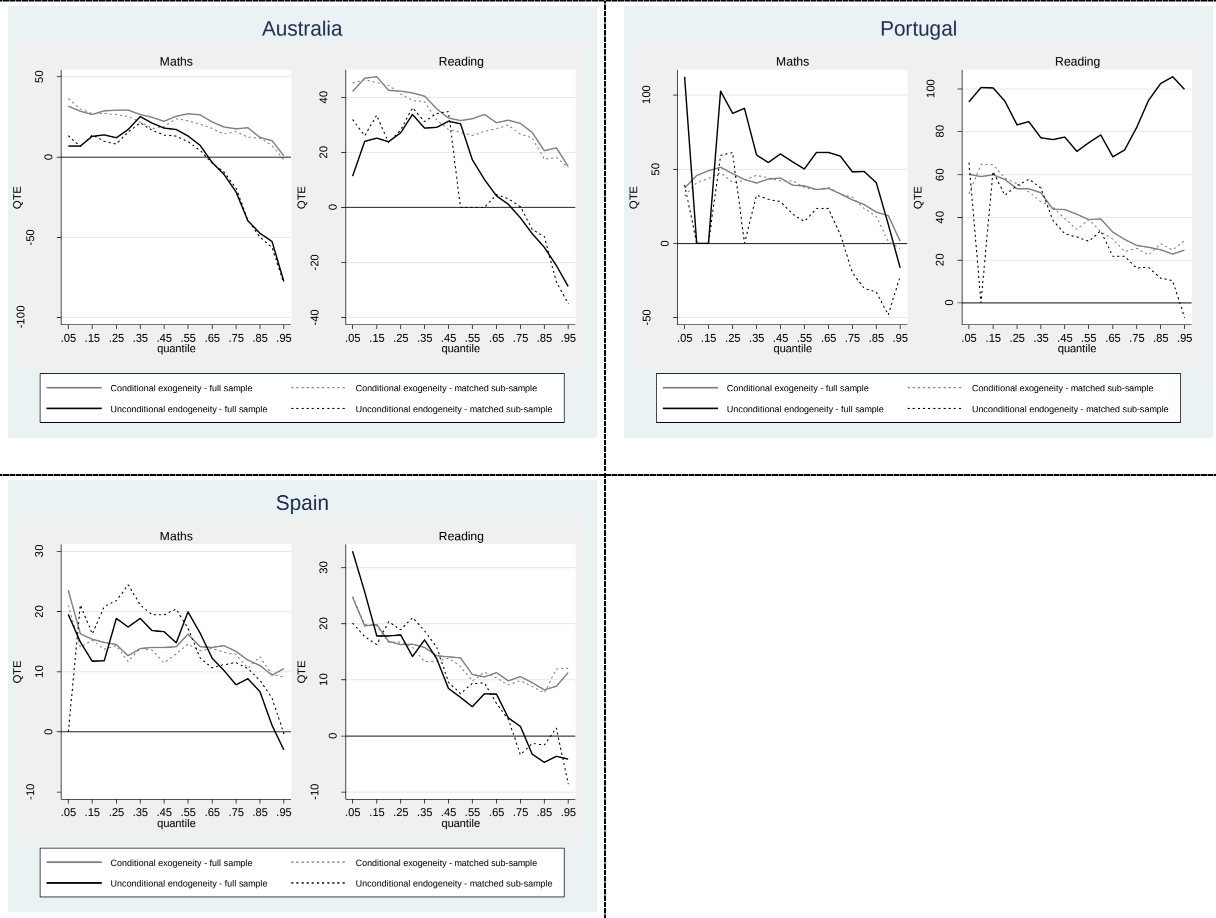
Notes: (1) Linear IV specification based on same controls as in Table 3. (2) The instrument variable z1 measures parents' endorsement of instruction, or religious philosophy when students are admitted at schools, and the instrument variable z2 is constructed as the interaction of z1 with a dummy for no residence conditions on school admission. (3) Endogeneity test is based on Sargan-Hansen statistics, the weak identification test is based on Kleibergen-Paap rk LM statistic, and the SWr2 denotes the Sanderson-Windmeijer (SW) first-stage F statistics.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2. *IV quantile treatment effects estimates for private school in full and matched sub-samples*

	Full sample			Matched sub-sample		
	Q(0.25)	Q(0.50)	Q(0.75)	Q(0.25)	Q(0.50)	Q(0.75)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Unconditional QTE						
Australia:						
Maths	30.92***	32.95***	16.98**	25.55***	25.94***	16.20**
Reading	45.39***	43.30***	28.41***	27.18**	36.39***	25.50***
N	1,954	1,954	1,954	1,392	1,392	1,392
Portugal:						
Maths	62.55***	45.10***	25.08***	53.44***	30.07***	21.42**
Reading	59.72***	33.92***	30.74***	57.04***	37.83***	27.91***
N	3,558	3,558	3,558	566	566	566
Spain:						
Maths	14.80***	14.64***	12.23***	16.20***	15.89***	13.71***
Reading	19.73***	13.28***	7.23**	18.68***	12.87***	7.63**
N	7,613	7,613	7,613	4,998	4,998	4,998
Panel B - IV unconditional QTE (based on z_1)						
Australia:						
Maths	12.07	17.14	-21.65	8.18	13.09	-19.24
Reading	28.21**	30.43***	-3.69	38.32***	n/a	0.37
N	1,936	1,936	1,936	1,379	1,379	1,379
Portugal:						
Maths	87.86**	55.08**	48.22	61.30**	19.86	-19.63
Reading	83.22**	70.85***	82.05***	54.88***	30.9	16.36
N	3,432	3,432	3,432	561	561	561
Spain:						
Maths	18.85**	14.80**	7.87	21.81***	20.41***	11.53**
Reading	18.04**	6.91	1.71	18.91***	7.55	-3.33
N	7,557	7,557	7,557	4,965	4,965	4,965
Panel C - IV unconditional QTE (based on z_2)						
Australia:						
Maths	31.24***	37.78***	19.63*	30.22**	40.74***	19.32*
Reading	38.32***	43.66***	30.34***	53.30***	47.475***	24.62**
N	1,936	1,936	1,936	1,379	1,379	1,379
Portugal:						
Maths	56.24	27.11	-23.13	12.07	-6.78	-50.86**
Reading	31.77	-2.97	-30.81	-16.68	-39.69	-100.32***
N	3,432	3,432	3,432	561	561	561
Spain:						
Maths	0.39	2.34	21.58	25.7	29.13	34.90**
Reading	-9.3	-1.97	-7.99	-1.44	-3.09	1.09
N	7,557	7,557	7,557	4,965	4,965	4,965

Notes: (1) See notes of Table 3. (2) Panel A estimates are unconditional exogenous QTE based on Firpo et al. (2007), and estimates of Panels B and C are unconditional QTE under endogeneity based on Frölich and Melly (2013). (3) The instrument variable z_1 measures parents' endorsement of instruction, or religious philosophy when students are admitted at schools, and the instrument variable z_2 is constructed as the interaction of z_1 with a dummy for no residence conditions on school admission. (4) The analysis is run using the Stata command ivqte (Frölich and Melly, 2010).
Significance levels: * $p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Figure B1. *Comparison of exogenous and endogenous QTEs of private school effects for full and matched subsamples based on z1 (religion instrument)*



NOTES

[1] For some recent research using TALIS, see for instance: Duyar et al. (2013) on teacher work attitudes; Gumus et al. (2013) on professional development; Vieluf et al (2013) on teacher self-efficacy; Sealy et al. (2016) on head masters school satisfaction.

[2] Another, albeit technical difference, is that by employing recent results on counterfactual decomposition (Chernozhukov et al., 2013) and their empirical application (Chen et al., 2016) we are able to derive confidence intervals estimates for each contributing factor to the learning gap (by using weighted bootstrapped).

[3] Note that how accountability is defined within PISA has been criticized as a technical concept with a comparative purpose of performance, thereby ignoring political accountability leading to lower trust within education systems (see, for instance: Ozga, 2013; Sahlberg, 2010).

[4] The TALIS-PISA 2013 link sample represents a 17%, 72%, and 35% of the PISA 2012 for Australia, Portugal, and Spain, respectively.

[5] Using matching techniques to obtain matched sub-samples to isolate observables has been used to study the private school effect (Chudgar and Quin, 2012) and also for the learning gap related to school violence (Delprato et al., 2017). A formal version of this approach is Coarsened Exact Matching (CEM) where the balance between the treated and control groups is chosen ex-ante by the research (see Iacus et al., 2011).

[6] These techniques are employed in empirical economics to study the wage gap (see, Fortin et al. 2011, and references therein).

[7] Unconfoundedness (or the conditional independence assumption [CIA]) states that assignment to treatment is independent of the outcomes, conditional on the covariates: $\left(\mathrm{Y}(1) - \mathrm{Y}(0)\right) \perp (\mathrm{S} = 1)|\mathrm{X}$. Unconfoundedness is untestable, although the TALIS-PISA link contains rich information on relevant covariates that describe the child, their home, and learning environments to model private school choice. The second assumption, overlap (or common support) condition, states that probability of assignment into the treatment is bounded away from zero and one: $0 < \Pr(\mathrm{S} = 1|\mathrm{X}) < 1$. This condition ensures that any combination of characteristics observed in the treatment group can also be observed among the control group. Appendix A contains a visual inspection of the validity of the overlapping assumption for our application.

[8] Nearest neighbour consists of an algorithm that matches each treated student with the non-treated student displaying the closest propensity score. The method is applied without replacement. The closeness of the propensity scores to find matches is defined by the value of the caliper which, following previous studies (e.g., Gou and Fraser, 2010), we set as $0.25 \times \sigma_{PS}$ (25% of the standard deviation of the estimated propensity score). We also employ radius and kernel matching for robustness (see Appendix A). The analysis is run using the psmatch2 Stata routine (Leuven and Sianesi, 2012).

[9] It should be noted that this empirical choice is because how the return of a characteristic varies by its differential relationship among public and private school students (i.e., the coefficient) would require further insights as it is not well understood in the literature in comparison to linking the effect to the characteristic's contextual or aggregate level.

[10] Specifically, Appendix A contains information on the quality of matching supporting our sub-sample matched analysis. Figures A1–A3 display the distribution of the propensity score and selected covariates in the full and matched samples by school type. These figures show that the gap in the probabilities of attending public and private schools is considerably narrowed in matched samples, whilst the distributions of household wealth and mother's education are similar by school type in the matched samples too. Table A1 further supports these results through mean tests on all students' covariates; we find no statistically significant difference in students' characteristics in public and private schools post-matching (i.e., all t critical values are below 1.96). Also, different matching procedures (nearest neighbor, radius and kernel) lead to similar estimates of the private effect, and the overall mean biases for all covariates are below 5% (see Table A2).